Exploring the Geographic Distribution of Patients Seen at the Columbus Free Clinic

Thesis

Presented in Partial Fulfillment of a Bachelor's of Science in Social Work in the College of Social Work at the Ohio State University

Cory J. Roth

Ohio State University

2017

Thesis Committee

Dr. Lisa Raiz, Advisor

Abstract

Currently, it is estimated that 13.3% of adults and 5.5 percent of children in the United States are uninsured. Uninsured or under-insured individuals report difficult in accessing care. While there are many factors that contribute to difficulty accessing healthcare, the largest barrier identified by patients is cost. Additionally, low income and uninsured individuals account for a disproportionate amount of healthcare morbidity and mortality in the United States. Community health centers and free clinics can act as a safety net, providing medical care to many individuals who may otherwise have to go to the emergency department or without treatment. This study was conducted at the Columbus Free Clinic (CFC), a clinic affiliated with the Ohio State University located in a large Midwestern city that provides medical, pharmacy, and social work services to patients free of charge. This study aims to understand the geographic distribution of the patient population of CFC against utilization, chronic condition, social work contact, and primary care cohort participation. A better understanding of the patient population is essential to offering more targeted, efficacious services to the patients of the clinic. Data were collected through chart review of patients seen at the clinic between January and June of 2016. Data were cleaned and analyzed at the zip code level. The data were visualized creating choropleth maps of the different variables. Cluster analysis was also used to propose significant subgroups within the larger population of the study. This clustering suggested between 4 and 6 distinct subgroups within the population based on patients and their utilization. While chronic condition, social work contact, and longitudinal patients had similar geographic distributions, the distribution for utilization information was significantly different, suggesting that some zip codes contribute many patients while others contribute patients who utilize the clinic at a higher rate. This information is useful in improving the clinics community based interventions and resource referrals.

Acknowledgments

I would like to start by thanking my parents and my siblings, for without them and their encouragement there would be no others to thank. I would also like to thank my Mentor, Dr. Raiz; because of the skills you have taught me and the value you have shown me I can produce; I plan to continue research as a part of my career. I would also like to thank Anna Stewart, as well as the board and steering committee of the Columbus Free Clinic. Not all individuals and organizations are as supportive as they have been, and this project is only possible because of them. I would like to thank Jennie Babcock and the Ohio State University College of Social Work for their support of my research efforts, as well as the professors who taught me in the foundational courses that prepared me for this study. I would like to thank the Ohio State University Department of Geography for teaching me programs and techniques that were vital to my study. Lastly, I would like to thank the patients of the Columbus Free Clinic. I hope this study serves as a token of my appreciation for all they have taught me.

Curriculum Vitae

May 2017……………………………B.S. Social Work, Honors with Research Distinction, Summa cum laude, The Ohio State University

Anticipated Graduation,

May 2021……………………………MD Candidate

The Ohio State University College of Medicine

Fields of Study

Undergraduate Field: Social Work

Graduate Field: Medicine

Table of Contents

Chapter 1: Statement of Research Topic

Introduction

In the United States, national healthcare expenditures reached $2.8 trillion, 17.2% of the nation's gross domestic product, far higher than any other developed country in the world in 2015 (Center for Medicaid and Medicare Services, 2015). In addition to these high healthcare cost, the United States has poor population health outcomes when compared to other nations (Askin & Moore, 2014). The disparities in population health are particularly marked among different geographic, economic, and demographic groups (Parrish, 2010) The United States center for Medicaid and Medicare Services (2015) projects healthcare spending to increase by about 6% each year. It has also been noted that this healthcare spending is not distributed equally among the population (Askin & Moore). Five percent of the individuals in the United States account for over 50% of the country's healthcare spending with the sickest 1% accounting for 21.8% of healthcare spending (Askin & Moore). There are numerous systemic and individual reasons that healthcare is so costly, however two of the most significant and interrelated reasons for high healthcare cost in the United States are lack of access to primary care and emergency department over-use (Douthit, Kiv, Dwolatsky, & Biswas, 2015; Harris, et al., 2016).

Access to Healthcare

It is estimated that 13.3% of adults and 5.5% percent of children are currently uninsured in the United States (Center for Disease Control and Prevention, 2014). Uninsured adults are more likely to report that they are not able to see physician when needed and had not had a regular checkup in the past 2 years (Ayanian, Weissman, Schnieder, Ginsburg, & Zaslavsky, 2000). The largest barrier to primary care identified by patients is cost (Kamimura, et al., 2015).

Studies have shown that access to primary care is associated with better health outcomes even when controlled for income. (Shi, Starfeild, Kennedy, & Kawachi, 1999; Shi & Starfeild, 2000). Additionally, low income and uninsured patients carry disproportionate health morbidity and mortality in the United States (Kaplan, Pamuk, Lynch, Cohen, & Balfour, 1996)

Use of Emergency Department

Many people who are uninsured or of lower income report few options in accessing medical treatment. Many individuals report that they can only access healthcare through the emergency department (ED) (Kamimura, et al., 2015). The groups with the highest ED usage rate tend to be uninsured or Medicaid recipients (Gindi, Black, & Cohen, 2016). These patients are often treated for chronic medical conditions which can require expensive ongoing treatment (Newton, Keirns, Cunningham, Hayward, & Stanley, 2008). The CDC reports that one third of ED visits are semi urgent or non-urgent and could be handled at other facilities

Free and Community Health Clinics

Because healthcare is a limited resource at any given time, the high costs the US healthcare system incurs have consequences in the way healthcare is rationed. One way the United States healthcare system is rationed is by restricting access based on ability to pay (Sommers, 2015). For this reason, among others, community health centers and free clinics act as a safety net, providing medical care to many individuals who may otherwise have to go to the ED or go without treatment (Posada, Potvin, & Kumar, 2014). Free clinics have been shown to be able to improve the unmet health needs of impoverished or uninsured individuals (Ayanian, Weissman, Schnieder, Ginsburg, & Zaslavsky, 2000; Ryskina, Meah, & Thomas, 2009). In 2014 the free clinics in the Ohio Association of Free Clinics (OAFC) provided $200 million worth of health care with $110 million used for diagnostic testing (Ohio Association of Free Clinics, 2014). Patients who utilize free clinics are typically low income, uninsured, disproportionately female, and nearly half come from rural areas, however, there is significant variation in patient population between free clinics (Notaro, et al., 2012). For this reason, it is important for each free clinic to conduct its own investigation of its patient population.

Statement of Problem

The Columbus Free Clinic (CFC) is a free healthcare clinic affiliated with the Ohio State University providing medical, pharmacy, and social work services to the Columbus and greater central Ohio area. Services provided include healthcare, lab testing, pharmacy, social work services, counseling, and referrals to specialty care. Anecdotal reports from the practitioners and providers at CFC have identified that the geographic distribution of the patient population is not well understood. Understanding the geographic distribution of CFC's patient population is essential to tailoring the services and resources provided in clinic. It has been shown that community based interventions can lead to higher quality of care and healthcare outcomes in both children and adults (Margolis, et al., 2005; McHugh, Harvey, Kang, Shi, & Scanlon, 2016).

Study Purpose

This exploratory study investigates the geographic distribution of patients seen at the Columbus Free Clinic between January and June of 2016. The purpose of this study is to describe the geographic distribution of the CFC patient population between January and June of 2016 investigating the following variables: zip code, diabetes, hypertension, hyperlipidemia, enrollment in clinic longitudinal primary care group, social work contact, and number of visits per six-month period. Analyzing the geographic distribution of the clinic's patients will help practitioners and providers in the clinic better understand the needs of the patient population's communities. This information will help CFC better tailor its service delivery to the needs of those communities and its patients. This research will help to inform future studies on the utilization behavior and environmental conditions of the clinic's patient population.

Chapter 2: Literature Review

Geographic Information Background

The effect of geography on health has been and continues to be a major topic of healthcare research (Basu & Siddiqi, 2014). New geographic information systems (GIS) and analytic techniques have also contributed to this field of healthcare research (Hu, Wang, Sun, Sorrentino, & Elbadollahi, 2012). Healthcare geographic studies mainly fall into three broad categories: establishing prevalence and incidence rates for diseases, assessing geographic accessibility and utilization, and identifying targets for resource intervention (Hu, Wang, Sun, Sorrentino, & Elbadollahi, 2012).

There is substantial evidence that understanding the geographic distribution of a patient population leads to more effective treatment (Hawthrone & Kwan, 2012). The first advantage of exploring the geographic distribution of a patient population is that it helps practitioners and providers identify targets of community based interventions that can improve health outcomes (Kaplan, Pamuk, Lynch, Cohen, & Balfour, 1996; Kamimura, et al., 2015; Hawthrone & Kwan, 2012). Many studies use geographic patient information and patterns of utilization to identify distinct groups within the patient population and to understand how resources should best be allocated (Abdullah, Laing, Hariri, Young, & Schafer, 2016; Muntner, et al., 2015). These studies often find areas of high utilization, and targeting these high utilization areas with resource interventions has been shown to improve healthcare in those areas (Harris, et al., 2016).

Subgroup Identification

A stated above, understanding the specific subgroups of a patient population can be useful in better addressing the needs of that patient population. A technique commonly used to understand different demographic and utilization trends within a geographic context is cluster analysis (Hu, Wang, Sun, Sorrentino, & Elbadollahi, 2012). Cluster analysis is an exploratory technique used to identify significantly similar groups within a larger population. Identifying these subgroups within the data set allows providers to tailor their interventions to different groups. This technique of targeting groups with specific interventions has been shown to be an effective intervention in many different contexts (Rodriguez, Wang, Naderi, Johnson, & Foody, 2013). On such context in which identifying subpopulations has been useful is within emergency departments. These practices have helped connect individuals with the resources they need and lower their ED utilization rate (Harris, et al., 2016). Particularly, this study by Harris et al was able to target individuals with difficulty managing chronic conditions and their overutilization of ED's. Understanding the geographic distribution of patients in this study allowed for more targeted community based interventions, a deeper understanding of environmental and health risk factors, and the resources available in a community. Understanding the geographic distribution has also been shown to elucidate the utilization patterns of different areas in this study (Harris, et al.). Understanding utilization patterns can be used to inform follow-up treatment and continuity of care, while simultaneously helping to identify barriers to access when understood in a geospatial context.

Another context in which these geographic techniques were employed was in a study by Rodriguez, Wang, Naderi, Johnson, and Foody (2013). Community risk factors for cardiovascular disease, including hypertension, and hyperlipidemia, helped to explain the prevalence of heart disease in different sub groups of the larger patient population. Many chronic condition, including hypertension, hyperlipidemia, and diabetes are heavily impact by individual's health behaviors (Fan, Strasser, Zhang, Fang, & Crawford, 2015; Gore & Kothari, 2013). Understanding the geographic distribution of patients in this study elucidated targets for environmental and community interventions for patients (Kamimura, et al., 2015). These interventions can help address both environmental and health behaviors risk factors for chronic conditions. Understanding a patient population's geographic distribution, as well as other demographic or chronic condition factors helps in identifying what resources are could be made available to patients in different areas to help them manage their chronic conditions and health (Rodriguez, Wang, Naderi, Johnson, & Foody). These resources can enhance services provided and compliment treatments in clinic, leading to overall better care.

Geographic Utility

Fisher and Skinner (2013) found there are wide geographic variation in per-individual healthcare spending and outcomes. These differences in the geographic distribution are impacted but not eliminated when controlling for confounding variables such as: age, sex, income, socioeconomic status, race, ethnicity, health status, or cost of services (Fischer & Skinner). This geographic effect, termed by Fischer and Skinner as unwarranted variations, ultimately suggests that healthcare spending in an area does not necessarily lead to better outcomes and that there are distinct effects in healthcare outcomes related to geographical location.

In addition to healthcare outcomes and spending, geographic information has also been used to describe healthcare accessibility in both rural and urban communities. Hawthrone and Kwan (2012) found that different populations face distinctly different geographic barriers to accessing their care. In their study, geography played a significant factor in perceptions of different healthcare providers, which informed healthcare access. In addition to geography, road networks and demographic factors played a significant role in the healthcare centers and services that individuals utilized found in this study.

For individuals who have difficulty accessing healthcare, community and free health clinics play an important role acting as a safety net (Ayanian, Weissman, Schnieder, Ginsburg, & Zaslavsky, 2000). While it is known that patients who utilize free clinics are typically low income, uninsured, disproportionately female, and nearly half come from rural areas, there is significant variation in patient population between free clinics (Notaro, et al., 2012). This literature review for this study found very few geographic studies concerning the geographic distribution of free clinics, and no studies that provide insight into central Ohio's free clinic utilizing populations.

Healthcare System Background

Because variations in healthcare spending and outcomes are linked to geography, it is important to understand the factors that contribute to healthcare spending and utilization. In 2012, the United States Spent 17.2% of its gross domestic product on healthcare, a total of $2.8 trillion or $8,423 per capita (Center for Medicaid and Medicare Services, 2015). For contrast, the median per capita cost of healthcare among 35 developed countries was $3199 in 2012 (Squires & Chloe, 2015). Since 2012 healthcare spending in the United States has increased to $9,451 per capita and is expected to continue to grow at a steady rate (Center for Medicaid and Medicare Services, 2015). Within the United States, healthcare spending can vary dramatically from state to state with Massachusetts having the highest healthcare spending of any state at $92,78 per capita, and Utah having the lowest healthcare spending of $5,031 per capita in 2012 (this does not include the District of Columbia which had a per capita spending of $10,349 during the same period) (Center for Medicaid and Medicare Services, 2015). Healthcare spending tends not be to distributed equally among the population either. In the US as a whole, 5% of the individuals in the account for over 50% of the country's healthcare spending with the 1% accounting for 21.8 percent of spending on healthcare (Agency for Healthcare Research and Quality, 2014). This national trend tends to also be represented at the state level where a relatively small number of individuals account for a majority of healthcare spending (Agency for Healthcare Research and Quality, 2014).

Factors Associated with Healthcare Spending and Outcomes

There are many documented and proposed causes of this high level of healthcare spending in the United States. One that has a direct impact on free clinic utilization is ED usage. ED utilization for non-emergent conditions contributes to rising healthcare costs (Center for Disease Control and Prevention, 2010). Under Current Laws, ED's must screen and stabilize all presenting patients regardless of ability to pay (Centers for Medicare and Medicaid Services, 2017). Groups with high ED utilization are Medicaid recipients, the poor, elderly individuals, and uninsured individuals (Center for Disease Control and Prevention). These groups are simultaneously more likely to have chronic conditions that require extensive, ongoing care and less likely to be able to pay for their treatment (Askin & Moore, 2014). Additionally, The CDC reports that one third of ED visits are semi urgent or non-urgent and could be handled at other facilities. The Free Clinic in which this study takes place receives patients referred from nearby EDs without the ability to access other sources of healthcare.

The Affordable Care Act tries to address many of these systemic issues and improve the United States Healthcare System including: expanding federal insurance programs, introducing accountable care organizations to lower readmittances and breaks in communication, eliminating insurer discrimination based on preexisting condition, and expanding mental health services.

Since the Affordable Care Act national insurance rates, both through employers and through government programs have increased without adversely impacting healthcare spending (Askin & Moore, 2014). This law is not perfect however; many argue that the exchange is inefficient and leads to higher premiums for individuals (Kocher, Emanuel, & DeParle, 2010). Issues of healthcare cost still exist since the implementation of this new law.

The consequences of these high healthcare costs are inequitable rationing of healthcare. Currently the US healthcare system rations based on restricting access to those with the ability to pay (Askin & Moore, 2014). Cost is commonly reported as the largest barrier for individuals to access healthcare (Sommers, 2015). Other barriers include: health literacy, social factors, public transportation systems, political systems, and physical geography(Sommers).

Chapter 3: Methodology

Study Design

This study is an exploratory study that utilizes secondary data obtained from the electronic health record (EHR) of a single free health clinic. The primary purpose is to investigate the geographic distribution and utilization of patients seen at the Columbus Free Clinic between January and June of 2016. The data used in this study were originally collected by practitioners and providers when patients were seen for treatment and medical care at the CFC. During patient visits, practitioners and providers encoded patient health information into the EHR, Practice Fusion, used by CFC. Practice Fusion is a free, fully HIPPA compliant, web based EHR.

Study Population

The population of this study is all patients seen at CFC from January to June of 2016. The patent population is adults as the clinic only serves adults. There is no charge for services at the clinic and no one is refused services. Additionally, language barriers are addressed through the use of translators and translating services. Between 20 and 30 patients are seen a week.

Variables

The variables of interest were zip code, diabetes, hypertension, hyperlipidemia, enrollment in clinic longitudinal primary care group, social work contact, and number of visits per six-month period. These variables were selected because both the existing literature and the clinic personnel indicate that better understanding the patient population could help inform the interventions provided and services rendered (Basu & Siddiqi, 2014). The sub group patients with diabetes, hypertension, hyperlipidemia, and social work contact were also investigated because management of these chronic conditions can be impacted by environment and health behaviors; and are often the target of community based interventions (Fan, Strasser, Zhang, Fang, & Crawford, 2015; Gore & Kothari, 2013; Hill, Nielson, & Fox, 2013). Anecdotal reports from the clinic's providers also suggest that hypertension, hyperlipidemia and diabetes are a common issue shared by many of the free clinics patients.

Data Collection

Zip-code, utilization rate, diabetes status, hypertension status, hyperlipidemia status, contact with social work services, and participation in clinic longitudinal primary care cohort information from January to June of 2016 was collected in this secondary analysis. Data were collected using internal features of the EHR used by CFC as well as the Remark Data collection system in order to develop the fullest picture of the patient population. The Remark Data Collection system is a tool where manual data coding sheets are created and read by a scanner into a computer. ICD 9 codes 401-405 were used to determine hypertensive status, ICD 9 codes 272 were used for hyperlipidemia status, and ICD 9 codes 249 and 250 were used to determine diabetes status. These codes were selected because they correspond to the chronic condition variables of interest. The ICD code system, short for the International Classification of Disease, is a classification system of all diseases that is used ubiquitously throughout medical record documentation and coding (World Health Organization, 2017). Medications coded into the medical record were also used as a supplement to ICD codes to determine chronic disease status in order to attempt to correct incomplete medical record coding. The six most popular medications of each disease using GoodRX popularity algorithm were included for each respective chronic condition. GoodRX is a website that aggregates prescription drug data from different pharmacies and allows users to identify the competitive cost of prescription medications (GoodRX, 2017). GoodRX's popularity ranking was used because GoodRX is a resource used by over 100,000 doctors in the United States, and regularly used by the CFC pharmacy to determine lowest out of pocket cost to patients (GoodRX, 2017). The medications used to determine hypertensive status were: Lisinopril, Amlodipine, Losartan, Metoprolol er, Furosemide, and Metoprolol. The medications used to determine hyperlipidemia status were: Atorvastatin, Simvastatin, Pravastatin, Rosuvastatin, Fenofibrate, and Lovastatin. The medications used to determine diabetes status were: Metformin, Metformin er, Lantus, Glipizide, Glimepiride, and Humalog. The presence of any of these medications in the medications list section of the EHR indicated that this patient would be included for the respective chronic condition variable. Other variables of interest, zip code, utilization rate calculated directly from the EHR, contact with social work, and participation in longitudinal primary care cohort were taken directly from the EHR and recorded using the Remark Data collection system. This information was hand coded onto the remark data collection form, and scanned using the Remark computer program. The Remark computer program scans and reads information from the hand coded data collection forms making data collection easier. A fidelity check was performed for the Remark data collection system. The first data collection sheet from each week of the study was verified against the Remark system's data entry for that sheet.

Data Cleaning

From the Remark computer program an excel worksheet and a comma separated values file were produced and read into the integrated development environment, R studio. In R studio, binary variables from the Remark data collection were reformatted and individual observations were grouped by zip code. This reformatted individual binary level data into zip code ordinal level data with the incidences being counted on a zip code level. For example, all the patient in the 43201 zip code with diabetes in their medical record were combined giving the total number of patients with diabetes in that zip code. Data cleaning was continued until each column represented one variable and each row represented one zip code. The initial data set contained 478 patients. One hundred and twenty-nine individuals were excluded from the data set due to not having geographic information in their medical record leaving 349 individuals. In addition to individuals excluded due to lacking geographic information, 5 outlier zip codes were excluded due to being located far from the central Ohio area. The outlier zip codes were: 17972, 44128, 44052, 45219, and 43608. Ultimately, a total of 52 zip codes and 349 unique patients were included in the study.

Most recent census information on zip code population was used to control for population differences among zip codes. For all variables except utilization, observed values were divided by zip code population totals and multiplied by 1000 (Hawthrone & Kwan, 2012). This was done to control for population differences. Utilization rates were divided by number of patients in the zip code to represent average patient utilization and control for skew in the utilization distribution (Hawthrone & Kwan, 2012). These controlled rates were then used in later visualization and analysis.

Data Analysis

This information was visualized through multiplayer choropleth maps made through the program ArcGIS. A choropleth map is a map that uses differences in shading or spacing a symbol to convey values over an area (Buckley, 2017). A k-means optimization test called the elbow method was coded using R Studio. K-means is an exploratory cluster analysis algorithm used to identify potential subgroups within a larger dataset (Lantz, 2015). The algorithm does this by maximizing the differences between groups of point data while minimizing the differences within groups. The script written for this study produces 19 cluster analyses, the first analysis used 2 groups and each subsequent analysis used included one more group until a total of 19 cluster analyses were produced. The heterogeneity, measured as the ratio of the sum of squares among clusters over the sum of squares within clusters, is graphed for each cluster. This method helps to identify the potential number of clusters within the data set. In order to prevent bias due to special autocorrelation zip code information was not included in the clustering algorithm.

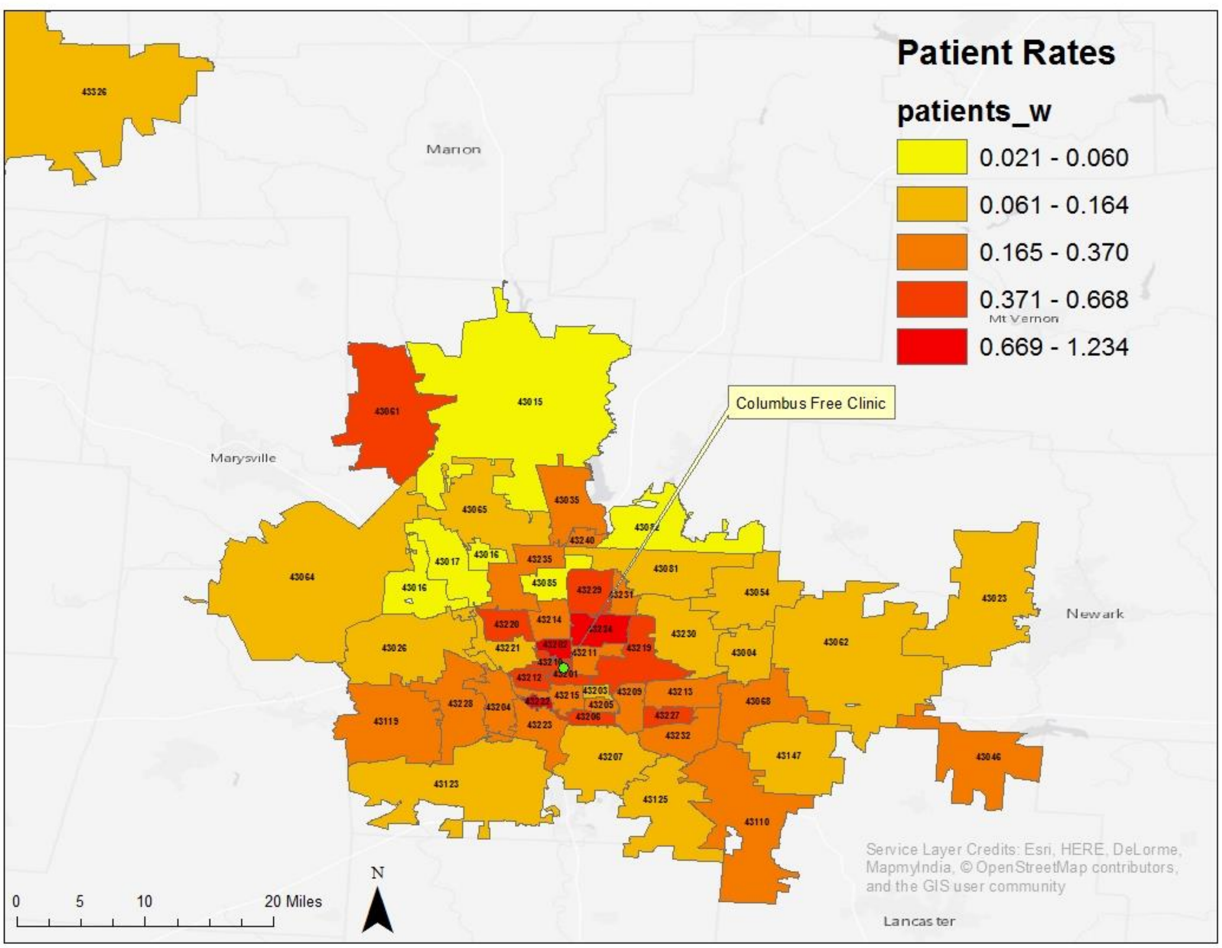
Chapter 4: Results

The 129 zip codes that were excluded from the study due to a lack of geographic information in the medical record can be considered either a limitation of the study or an interesting result. The same principle applies to the 5 outlier zip codes that were excluded from the analysis of this study. These outliers and individuals without zip code information will further be discussed in chapter 5.

Weighted Patient Contribution Rates by Zip Code

Based on the visualization, most highly contributing zip codes appear to be in a near proximity to the clinic. The highest quintile contributing zip codes were 43202, 43224, and 43222. These zip codes contributed patients at a rate of between 1.234 and 0.669. While CFC is not located in any one of these zip codes, these zip code areas are located near the clinic. The next quintile also appears to be located near the Columbus area. The first and second quintile appear to be contained within the 270-interstate belt that surrounds the Columbus area. These zip codes are: 43227, 43206, 43212, 43201 43210, 43219, 43220, and 43229. This second quintile contains the zip code in which CFC is located, 43201, as well. This quintile of zip codes contributed patients at a rate of between 0.668 and 0.371. The middle quintile was more broadly distributed in the Columbus and central Ohio area. While the majority of these zip codes are located south of CFC, some were located north of the clinic and these zip codes also appear to be distributed both west and east of the clinic. These zip codes contributed patients at a rate of 0.370 and 0.165. This quintile includes: 43119, 43228, 43204, 43223, 43215, 43211, 43205, 43209, 43213, 43068, 43232, 43110, 43046, 43214, 43235, 43240, 43035, and 43326. This middle quintile zip code displays considerable geographic heterogeneity. This zip code contains both urban areas, such as downtown Columbus, and rural mid-Ohio areas. The second to lowest contributing quintile was generally distributed around the peripheries of the catchment area and all but one zip code (43221) are located outside the 270-interstate belt. These zip codes contributed at a rate of between 0.164 and 0.061 and included: 43123, 43207, 43125, 43147, 43023, 43062, 43004, 43054, 43230, 43081, 43221, 43026, 43064, and 43065. The lowest contributing quintile is made up of five zip codes, 43016, 43017, 43015, 43085, and 43082, which contributed patients at a rate of between 0.060 and 0.021. All zip codes in this lowest quintile except 43085 are located outside the 270-interstate belt and these zip codes are also all located north of CFC.

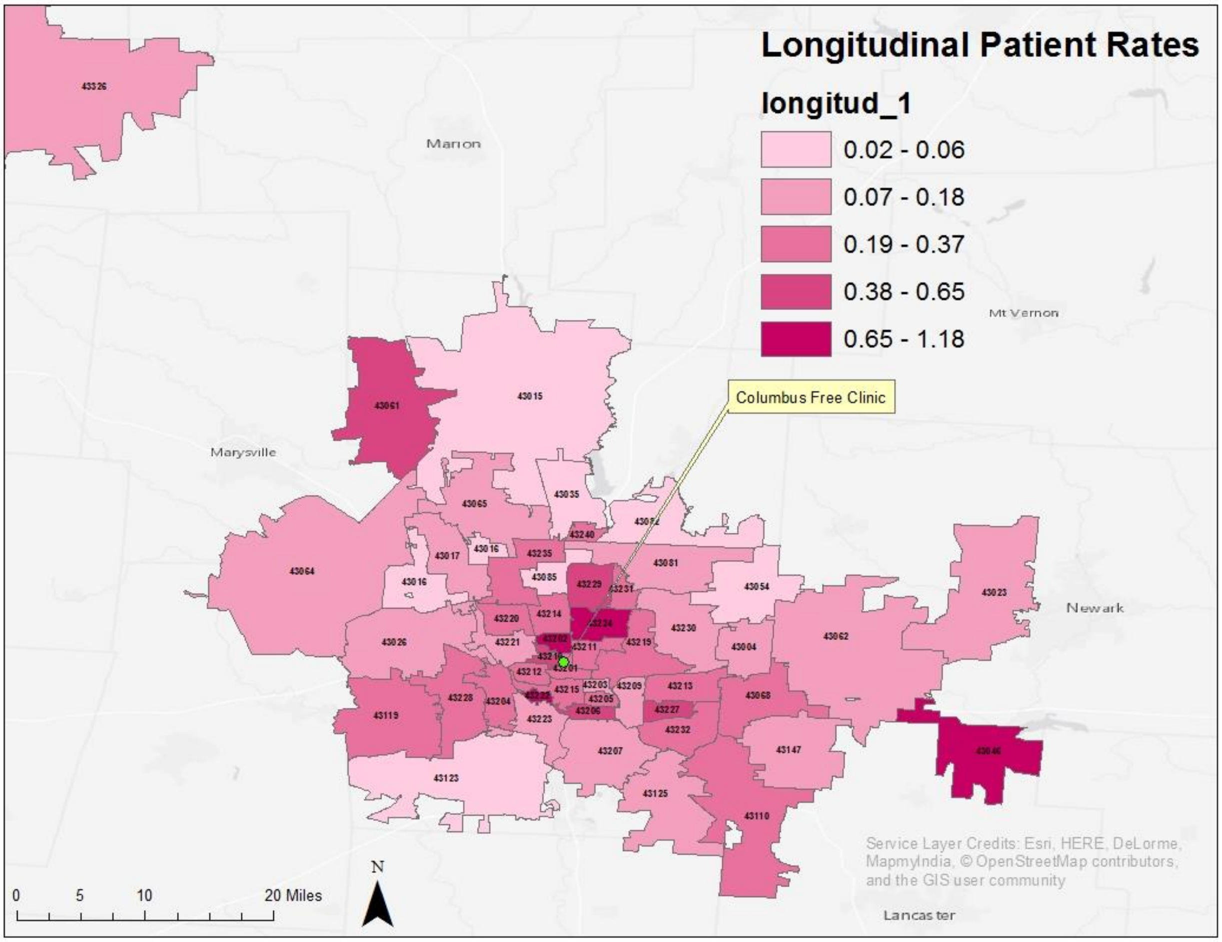
Figure 1. Patient Contribution Rate by Zip Code



Weighted Longitudinal Patients

The distribution of longitudinal patients shares some features of the general patient distribution; however, it also contains some significant differences. The zip code with the largest weighted amount of longitudinal patients controlling for zip code population contains 43046, 43222, 43202, and 43224. This quintile contributed between 1.16 and 0.65 percent of its population as longitudinal patients. While generally located near the clinic, 43064 is located on the far eastern side of the catchment area. The second quintile is made up of one zip code, 43061, that contributed 0.50 percent of its population as longitudinal patients. The third quintile contains the following zip codes: 43119, 43228, 43204, 43210, 43206, 43227, 43229, and 43240 and contributes longitudinal patients at a rate between 0.37 and 0.19. the second to last quintile contributes longitudinal patients at a rate between 0.18 and 0.07. the fourth quintile contains the zip codes: 43064, 43026, 43221, 43017, 43065, 43223, 43207, 43215, 43147, 43023, 43062, 43004, 43230, 43081, and 43326. This quintile appears to be distributed widely over the central Ohio catchment area. The final quintile has a longitudinal patient contribution rate between 0.06 and 0.02 and included: 43123, 43054, 43085, 43016, and 43015. All but one of these zip codes (43123) are located north of CFC and all are located outside the 270-interstate belt.

Figure 2. Longitudinal Patient Rate by Zip Code



Chronic Conditions

Disease codes for hypertension, hyperlipidemia, and diabetes were collected in this study. The geographic distribution for these three conditions appeared to be very similar with hypertension and hyperlipidemia presenting slightly more often. Hypertension and hyperlipidemia had exactly the same distribution while the diabetes distribution was slightly different. Four zip codes were in the highest quintile across all three conditions. These zip codes were 43229, 43224, 43202, 43210. In the diabetes category, 43222 and 43227 were also in the highest quintile. The second quintile in each category was made predominately of zip codes centering around CFC. These zip codes are approximated by the 270-interstate. The third quantile across the conditions in generally located south of the clinic. These zip codes include: 43119, 43228, 43204, 43223, 43209, 43068, 43125, 43110, and 43124. For hypertension and hyperlipidemia 43235 and 43231 were also third quintile zip codes. For hyperlipidemia and hypertension, the fourth quintile was made up of 43326, 43064, 43026, 43035, 43081, 43230, 43203, 43062, 43147, 43207, 43125, and 43023. This is nearly in identical to the diabetes fourth quintile distribution. Most these zip codes are located south of the free clinic. For hypertension and hyperlipidemia, the firth quintile contained the zip codes: 43123, 43004, 43054, 43016, 43017, 43015, 43065, 43085, and 43082. For diabetes, the lowest quintile contained the zip codes: 43147, 43062, 43054, 43082, 43231, 43085, 43016, 43065, and 43015.

Figure 3. Diabetic Patient Rate by Zip Code

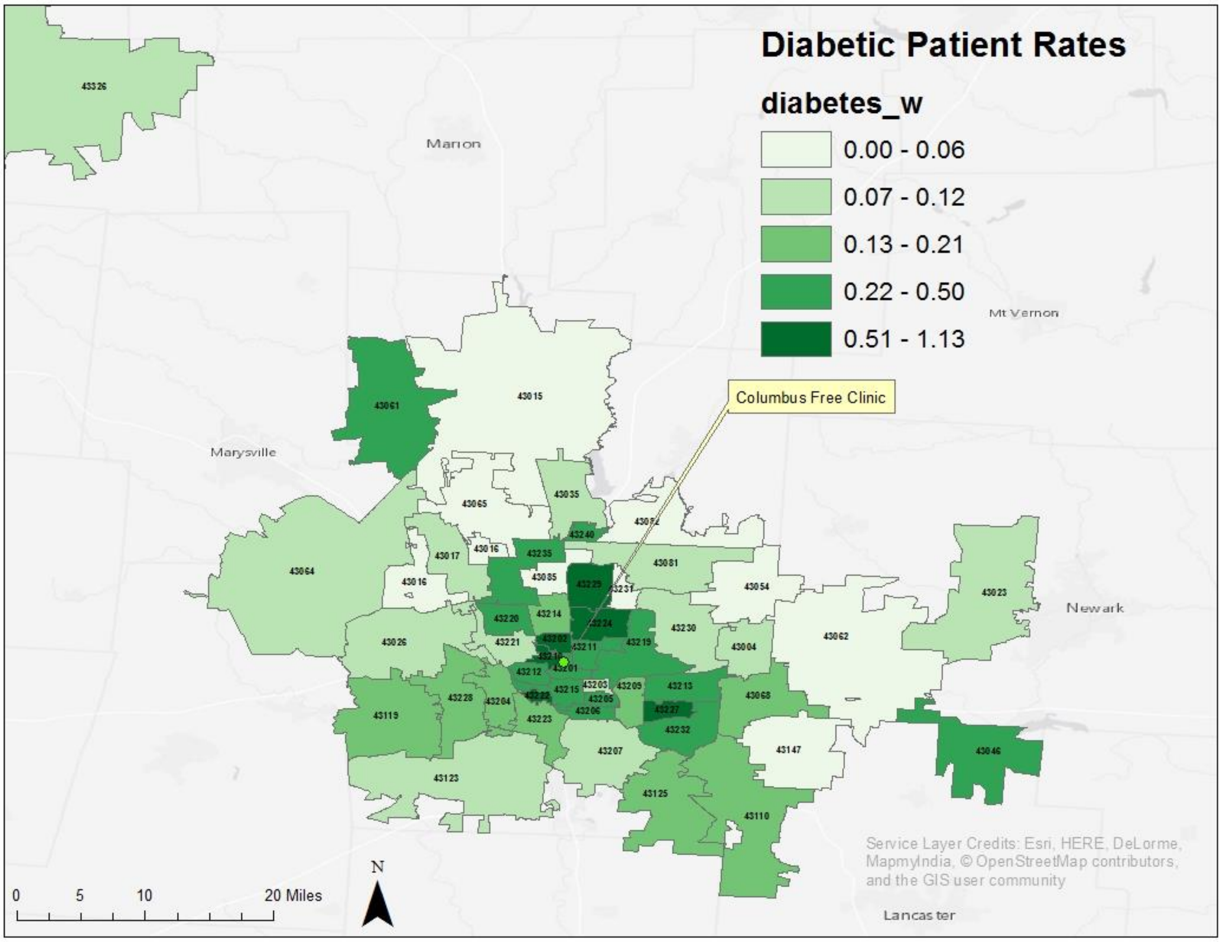
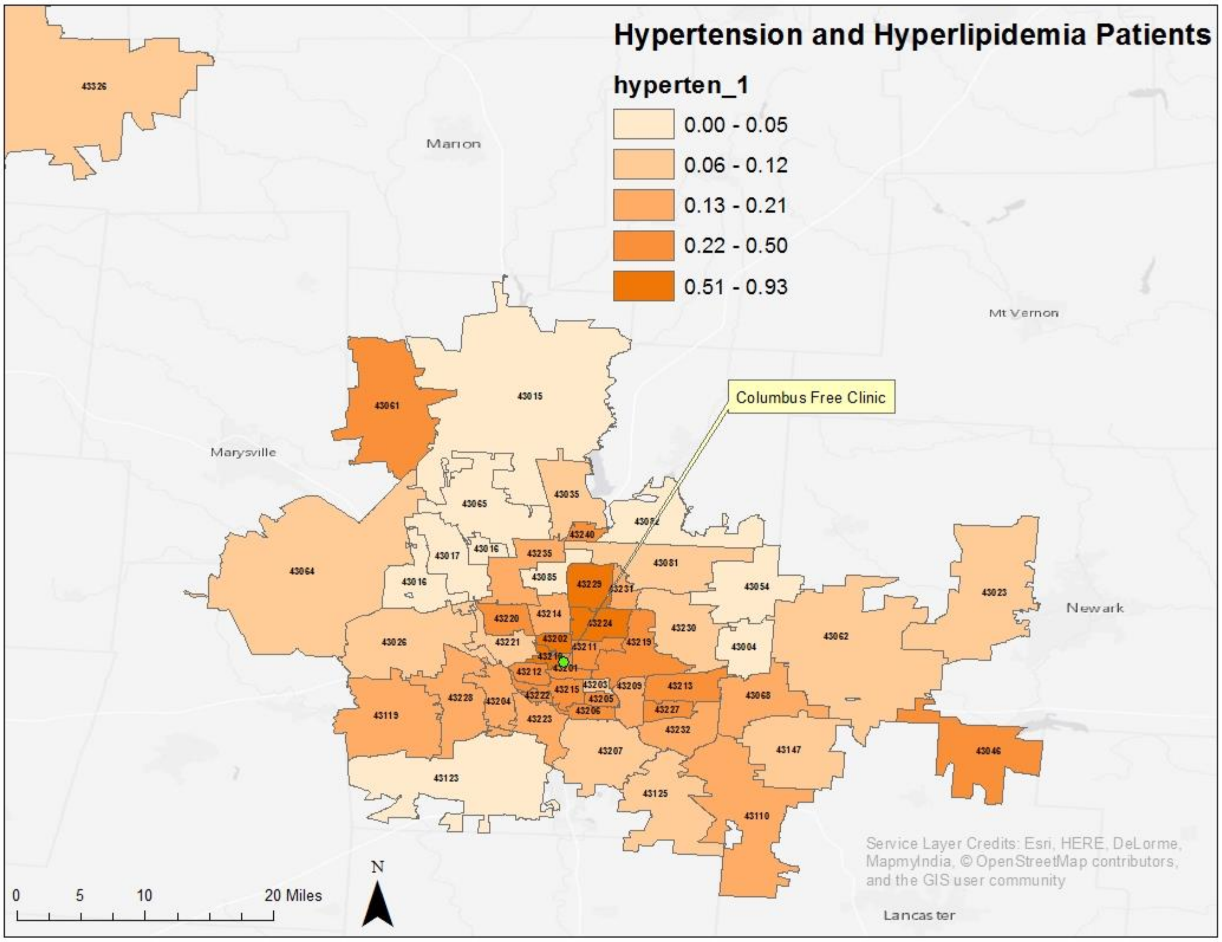


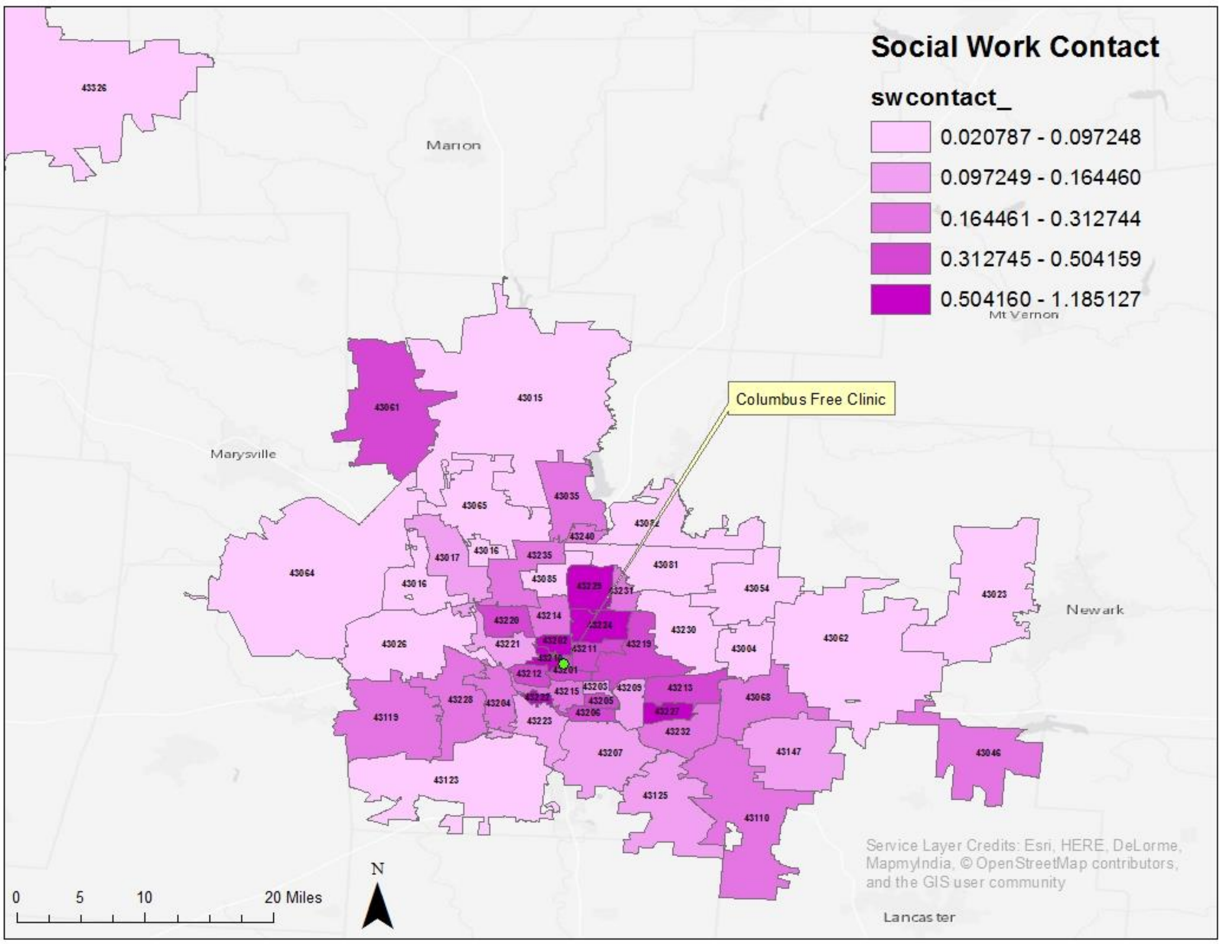
Figure 4. Hypertension and Hyperlipidemia Patient Rate by Zip Code



Social Work Utilization

There were seven zip codes in the upper quantile for social work utilization: 43061, 43229, 43224, 43202, 43210, 43222, and 43227. These zip codes are generally centered around the clinic with all but 43227 and 43061 being located inside the 270-interstate ring. The second quintile located in proximity to the clinic and concentrated inside the I-270 landmark. This quintile includes 43206, 43205, 43212, 43201, 43211, 43219, and 43220. The third quintile in spread more profusely throughout the southern portion of the catchment area. The zip code contained in the third quintile were 43119, 43228, 43204, 43232, 43068, 43110, 43046, 43214, 43235, 43240, 43035. The fourth quintile was made up of zip codes: 43017, 43221, and 43202; and the final quintile of social work utilization was made up of 43326, 43015, 43065, 43016, 43085, 43026, 43064, 43123, 43023, 43026, 43004, 43230, 43054, 43081, and 43082. These zip codes are generally distributed throughout the catchment area.

Figure 5. Social Work Contact Rate by Zip Code

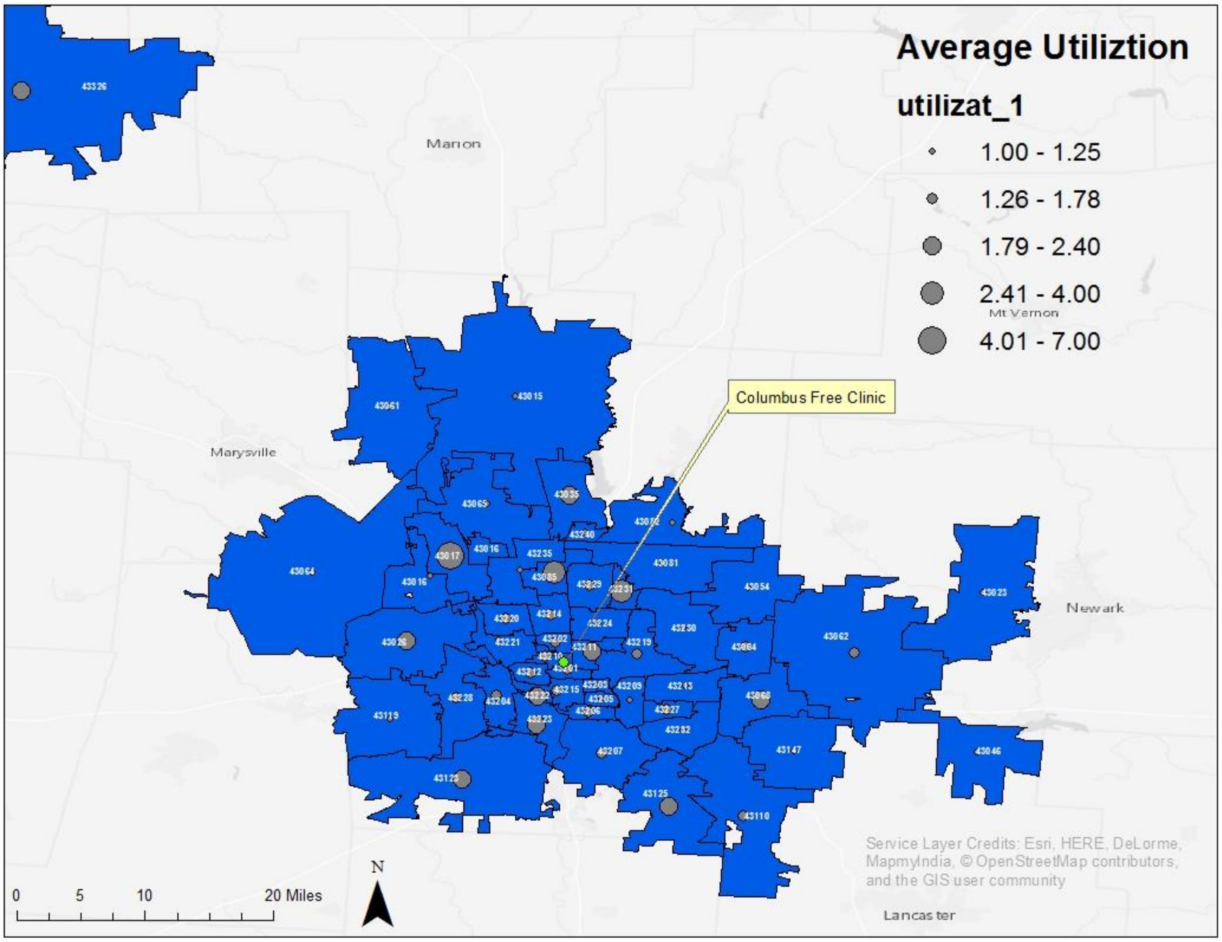


Utilization of Clinic

Utilization was visualized as average utilization of those who visited the clinic from each zip code. The first grouping was an average utilization rate of 1.00-1.25 visits during the time-period of the study. The second quintile had a utilization of between 1.26 and 1.78 during the study. The third quintile had a rate between 1.79 and 2.40 visits on average. The fourth quintile had an average utilization between 2.41 and 4.00 while the final quintile had an average utilization of 4.01 and 7.00. Higher Utilizing patients appear to be concentrated more centrally, however there are also some significantly utilizing groups on the farther reaches of the catchment area of the clinic.

When weighted patient contribution was overlaid by average utilization data distinct interesting groups emerged. High contributing zip codes tended to be low utilizing and areas with high average utilization tended to contribute a low number of patients. The k-means optimization showed that there were possibly 4-6 verifiable distinct groups within the data set. Heterogeneity of the groups decreased until the 4-6 cluster range, at which time heterogeneity leveled off. This point of inflection indicates that there are likely distinct subgroups among these variables within the larger distribution.

Figure 6. Average Patient Utilization by Zip Code



Utilization and Number of Patients

Patient utilization was visualized in addition to patient contribution rates of zip codes. Many of the zip codes that contributed many patients per capita were also the zip codes that were found to have the higher numbers of longitudinal patients, contact with social work, and chronic conditions. The zip codes that were consistently in the top quintile for social work contact chronic conditions and number of patients were 43202, 43224, and 43223. This pattern did not sustain for utilization information. The zip codes that were in the highest quintile for utilization were 43217, 43085, 43231, 43222, and 43223. The k-means optimization shows evidence for between 4 and 6 significant subgroups across number of patients and patient utilization as the graph inflects between 4 and 6 unique clusters.

Figure 7. Patient Contribution Rate and Average Utilization by Zip Code

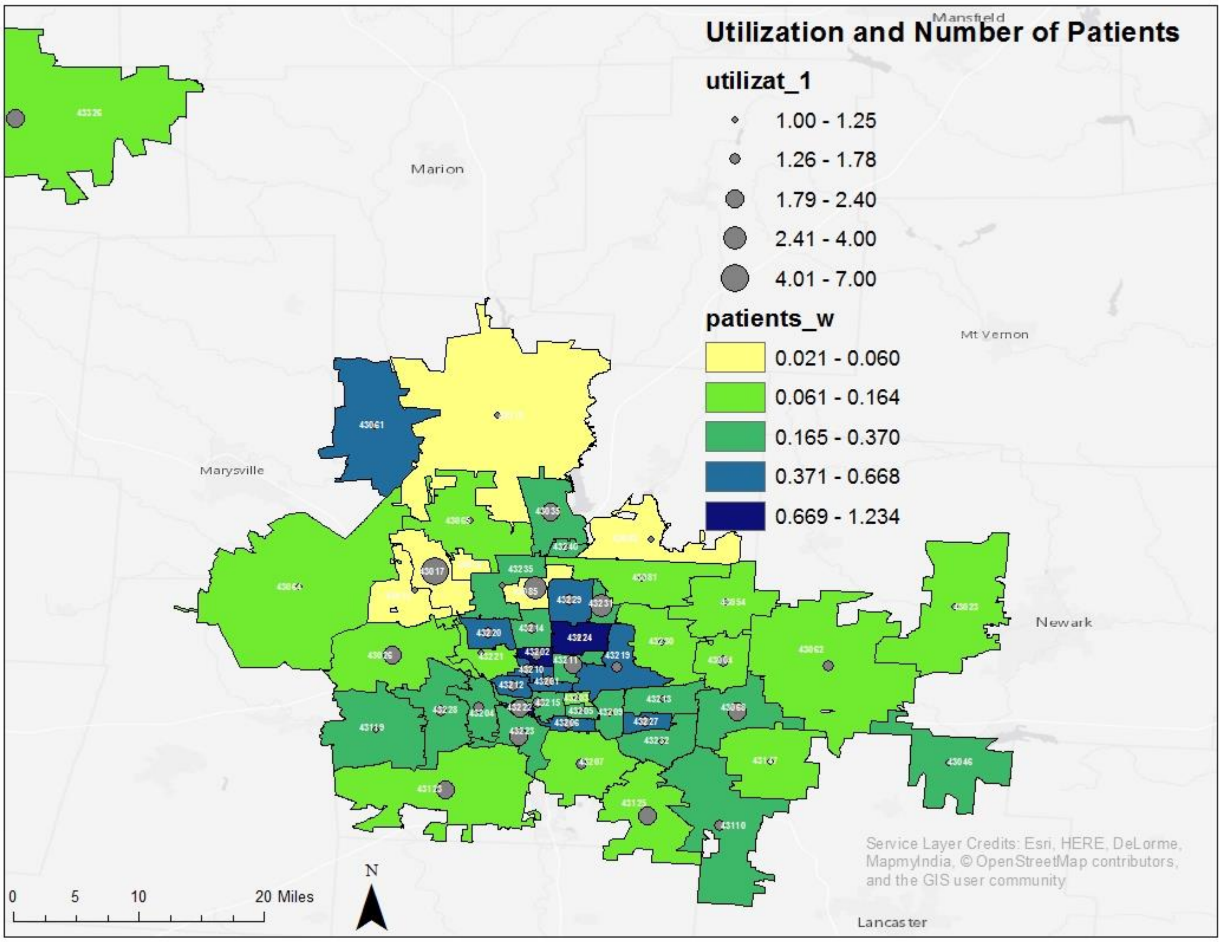
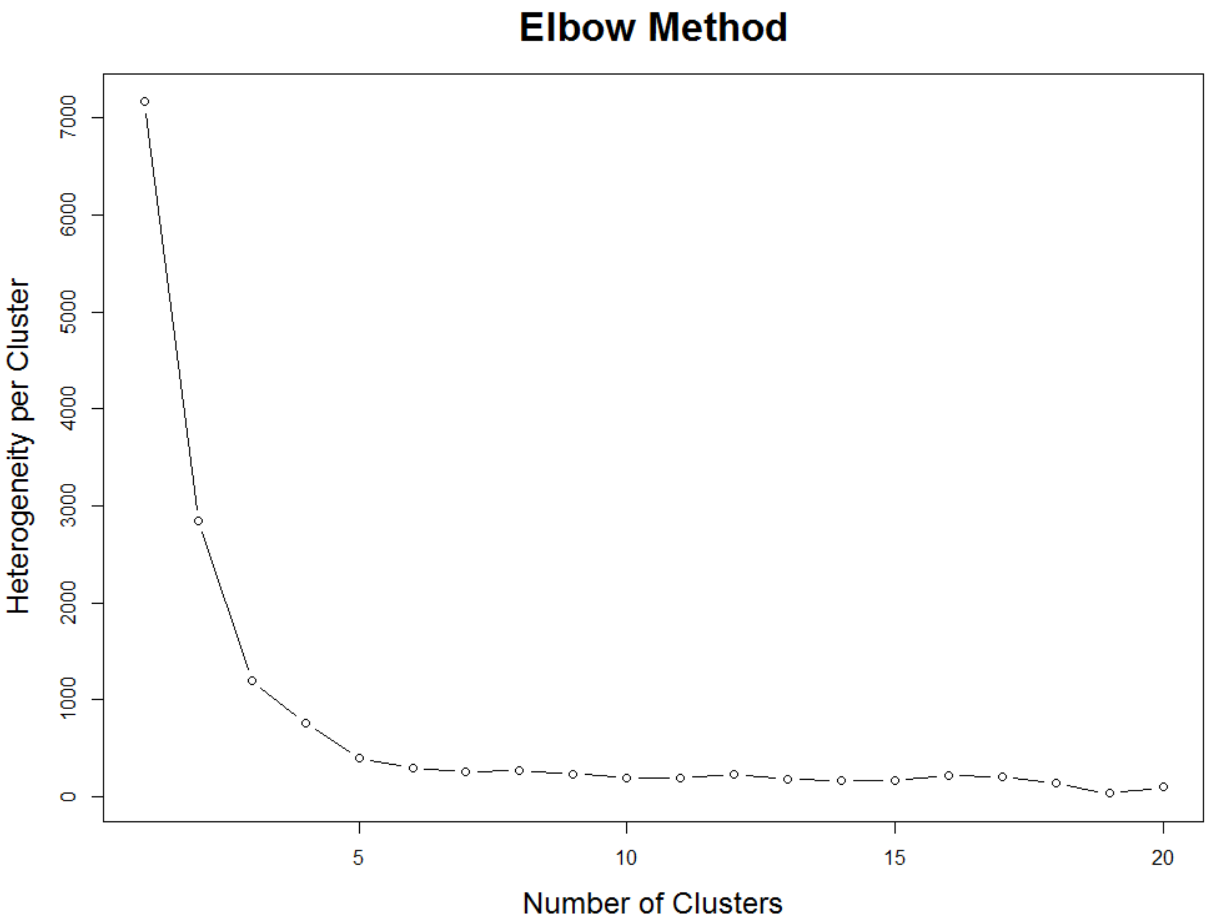


Figure 8. Elbow Method K-Means Optimization



Chapter 5: Discussion

Summary of Results

This study found a group of zip codes that were regularly in the highest quintile of each of the variables of interest. These zip codes tended to be located within the I-270 belt in proximity to the clinic relative to the larger catchment area. There also appeared to be a grouping of zip codes that were constantly in the lower quintile of the variables of interest. These zip codes appeared to be located north of the clinic outside of the Interstate 270 belt. Within the context of these results there were also some interesting zip codes that appeared not to adhere to general trends, either because of the zip code's patient utilization behaviors or because of their geographic location. Combined utilization behavior also presented interesting findings as well. When weighted patient contribution was overlaid by average utilization data distinct interesting groups emerged. High contributing zip codes tended to be low utilizing and area with high average utilization tended to contribute a low number of patients. The k-means optimization showed that there were possibly 4-6 verifiable distinct groups within the data set.

Limitations

The sample for this study is a convenience sample, and not generalizable to individuals outside of the sample. Additionally, many patients (129) did not report zip code information and subsequently their information was excluded from the data set. In one light this is a limitation of the study, but in another light this is in it of itself an interesting finding. The fact that 129 patients did not have geographic information recorded in their medical record either sheds information on the free clinic's population, the way in which providers at the clinic interact with patients or both. Additionally, 5 zip codes from outside the mid-Ohio area where treated as outliers in this study. While a limitation that these patients could not be included in this analysis, there five zip codes represent an interesting finding and should be considered separately. Lastly, Instances in which information was incorrectly entered into the medical record could not be accounted for in the secondary data collection.

Practice Implications

The results of this study have many significant implications for the clinic's personnel, policies, and procedures. While called the Columbus Free Clinic, this study shows that the patient population of the clinic come from many different parts of the central Ohio area. Providers and practitioners at the clinic should be aware of this larger catchment area and patient distribution as this information should inform the community based interventions provided in clinic. Evidence suggests that the interventions provided to patients in clinic will be more efficacious if the referrals and resources for other services or benefits consider the geographic location of the patients (Kaplan, Pamuk, Lynch, Cohen, & Balfour, 1996; Kamimura, et al., 2015; Hawthrone & Kwan, 2012). Currently, some of the resources we use are targeted to the Columbus area. Expanding these interventions to also cover rural or mid-Ohio areas would allow CFC to better serve our patients.

Additionally, it was believed that many of our patients came from impoverished or disadvantaged areas within the mid-Ohio area. While some of the zip codes and patients served do come from zip codes with high poverty levels, it is also clear that there are many patients who also come from areas with far below average poverty rates. It is very easy to build a stereotypical conception of the patients seen at the Columbus free clinic that may not represent of all patients. Providers should consider that CFC serves patients from many different areas and backgrounds and try to suspend or combat absolute or preconceived notions about the patient population.

It should also be taken into account that different patients have different utilization behaviors. While anecdotally, care coordination and follow up is believed to be a significant problem, this study shows that there are some patients that utilize the clinic sparingly and some that might be considered super utilizers in the distribution. The course of treatment for a patient who is unlikely to return to the clinic may need to be very different from the course of treatment for an individual that heavily utilizes the clinic. Utilization pattern can have implications for everything from resource acquisition for patients, to monitoring of prescription medications.

Future Research

The findings of this study indicate a need to further understand the CFC patient population in order to ultimately provide better services. Zip codes are considerably heterogeneous, and contain many different communities and neighborhoods. Analyzing the distribution of patients on a more granular level may elucidate some nuanced utilization patterns. For example, looking at the geographic information on the level of point data rather than zip code data could show areas in which patients are concentrated within zip codes. Additionally, many more chronic conditions can be geographically described in order to better understand the patient population. Different clustering analyses could be run in order to take into account geographic data and try to make meaningful groups using location without introducing bias.

Conclusion

The finding that the distribution of utilization behaviors was significantly different from other variables is interesting. It indicates that significant groups may exist within the larger patient population of the study. These populations should be further explored and validated with additional analyses. Understanding these subgroups based on utilization behavior and geography could allow the clinic to provide more targeted interventions to the population.

References

Abdullah, N., Laing, S., Hariri, S., Young, C., & Schafer, S. (2016). Use of claims data to estimate annual cervical cancer screening percentages in Portland metropolitan area, Oregon. Cancer Epidemiology, 41(1), 106-112.

Agency for Healthcare Research and Quality. (2014). The Concentration and Persistence of in the Level of Healthcare Expenditures Over Time. Retrieved from Agency for Healthcare Research and Quality: https://meps.ahrq.gov/data_files/publications/st354/stat354.pdf

Askin, E., & Moore, N. (2014). The Healthcare Handbook: A Clear and Concise Guide to the United States Health Care System. St. Louis: Washington University in St. Louis.

Ayanian, J., Weissman, J., Schnieder, E., Ginsburg, J., & Zaslavsky, A. (2000). Unmet Health Needs of Uninsured Adults in the United States. Journal of the American Medical Association, 248(16), 2061-2069.

Basu, S., & Siddiqi, A. (2014). Geographic Disparities In US Mortality "Hot Spotting" Large Data Bases. Epidemiology, 25(3), 467-470.

Buckley, A. (2017). Arc GIS Blog. Retrieved from ESRI.

Center for Disease Control and Prevention. (2010). Emergency Department Visitors and Visits: Who Used the Emergency Room in 2007? Atlanta: Center for Disease Control and Prevention.

Center for Medicaid and Medicare Services. (2015). NHE Fact Sheet. Retrieved from Center for Medicaid and Medicare: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/

Centers for Medicare and Medicaid Services. (2017). Emergency Medical Treatment & Labor Act (EMTALA). Retrieved from CMS.gov: https://www.cms.gov/regulations-and-guidance/legislation/emtala/

Douthit, N., Kiv, S., Dwolatsky, T., & Biswas, S. (2015). Exposing some important barriers to health care access in the rural USA. Public Health, 129(1), 611-620.

Fan, A., Strasser, S., Zhang, X., Fang, J., & Crawford, C. (2015). State Socioeconomic Indicators and Self Reported Hypertension Among US Adults, 2011 Behavioral Risk Factor Surveillance System. Preventing Chronic Disease, 12, 27-32.

Fischer, E., & Skinner, J. (2013). Making sense of geographic variations in health care: the new IOM report. The Journal of Health Affaris, 4(1), 203-210.

Gindi, R., Black, L., & Cohen, R. (2016). National Health Statistics Reports Reasons for Emergency Room Use Among U.S. Adults Aged 18–64: National Health Interview Survey, 2013 and 2014 . Atlanta: Center for Disease Cotnrol and Prevention.

GoodRX. (2017). How GoodRX Works. Retrieved from GoodRX: https://www.goodrx.com/how-goodrx-works

Gore, D., & Kothari, A. (2013). Getting to the Root of the Problem: Health Promotion Strategies to Address the Social Determinants of Health. Canadian Journal of Public Health, 104(1), 52-54.

Harris, J., Graetz, I., Podila, P., Wan, J., Waters, T., & Bailey, J. (2016). Charecteristics of Hospital Emergency Care Super-Utilizers with Multiple Chronic Conditions. Journal of Emergency Medicine, 50(4), 203-214.

Hawthrone, T., & Kwan, M.-P. (2012). Using GIS and perceived distance to understand the unequal geographies of healthcare in lower-income urban neighbourhoodsgeoj_411 18.30. The Geographical Journal, 178(1), 18-30. doi:10.1111/j.1475-4959.2011.00411.x

Hill, J., Nielson, M., & Fox, M. (2013). Understanding the Social Factors That Contribute to Diabetes: A Means to Informing Health Care and Social Policies for the Chronically Ill. Permanente Journal, 17(2), 67-72.

Hu, J., Wang, F., Sun, J., Sorrentino, R., & Elbadollahi, S. (2012). A Healthcare Utilization Analysis Framework for Hot Spotting and Contextual Anomaly Detection. American Medical Informatics Assocation Journal, 360-370.

Kamimura, A., Ashby, J., Jess, A., Trinh, H., Nourian, M., Finlayson, S., . . . Reel, J. (2015). Impact of Neighborhood Enviroments on Health Consciousnness, Information Seeking, and Attitudes US-Born and non US_Born Free Clinic Patients. Southern Medical Journal, 108(12), 703-709.

Kaplan, G., Pamuk, E., Lynch, J., Cohen, R., & Balfour, J. (1996). Inequality in Income and Mortlity In the United States: Analysis of Mortality and Potential Pathways. British Medical Journal, 312(7037), 999-1003.

Kinchen, K., & James, W. (1991). Hypertension Management in Health Care for the Homeless Clinics: Results from a Survey. American Journal of Publi Health, 81(9), 1163-1165.

Kocher, R., Emanuel, E., & DeParle, N.-A. (2010). The Affordable Care Act and the Future of Clinical Medicine: The Opportunities and Challenges. Annals of Internal Medicine, 153(8), 536-539.

Lantz, B. (2015). Machine Learning With R. In B. Lantz, Machine Learning With R (2nd ed., pp. 285-310). Birmingham: Packt Publishing Ltd.

Moore, W., & al, e. (2003). Medical Errors Related to Disconteutiy of Care from Inpatient to Outpatient Setting. Journl of General Internal Medicine, 18(8), 645-659.

Muntner, P., Gutierrez, O., Zhao, H., Foz, C., Wright, N., & Curtis, J. (2015). Validation Study of Medicare Claims to Identify Older US Adults With CKD Using the Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study. American Journal of Kidney Disease, 65(2), 249-258.

Newton, M., Keirns, C., Cunningham, R., Hayward, R., & Stanley, R. (2008). Uninsured Adults Presenting to Emergency Departments Assumptions vs Data. Journal of the American Medical Association, 300(16), 1914-1924.

Parrish, G. (2010). Measuring Population Health Outcomes. Preventing Chronic Disease, 7(4), A71.

Posada, J., Potvin, H., & Kumar, A. (2014). How a Student Run Clinic Can Address Gaps in Access to Mental Health Services. Texas Public Health Association Journal, 64(4), 9-12.

Rodriguez, F., Wang, Y., Naderi, S., Johnson, C., & Foody, J. (2013). Community-Level Cardiovascular Risk Factors Impact Geographic Variation in Cardiovascular Disease Hospitalizations for Women. Journal of Community Health, 38(1), 451-457. doi:10.1007/s10900-012-9640-2

Sommers, B. (2015). Health Care Reform's Unfinished Work — Remaining Barriers to Coverage and Access. The New England Journal of Medicine, 373(25), 2395-2397. doi:10.1056/NEJMp1509462

Squires, D., & Chloe, A. (2015). U.S. Health Care from a Global Perspective Spending, Use of Services, Prices, and Health in 13 Countries. The Common Wealth Fund.

Tao, L., Liu, J., & Xiao, B. (2013). BMC Health Services Research. Retrieved from BioMed Central.

Terris, M. (1996). The Development and Prevention of Cardiovascular Disease Risk Factors: Socioenvironmental Influences. Journal of Public Health, 17(4), 426-441.

Watson, J. (2014). Young African American Males: Barriers to Access to Health Care. Journal of Human Behavior in the Social Environment, 24(1), 004–1009.
doi:10.1080/10911359.2014.953416

World Health Organization. (2017). Classification. Retrieved from World Health Organization: http://www.who.int/classifications/icd/en/

Appendix A: Letter of Support From Columbus Free Clinic Board

COLUMBUS
FREE CLINIC

November 3, 2016

Dear Behavioral Institutional Review Board,

As a representative of the Columbus Free Clinic, I am writing on behalf of the Columbus Free Clinic steering committee to express our support for the proposed study "Geographic Distribution of Patients Seen by the Columbus Free Clinic" number 2016B0415 by principle investigator Dr. Lisa Raiz and co-investigator Cory Roth.

The Columbus Free clinic seeks to provide patients with a range of different health services. We employ community-based interventions in order to improve the quality of services provided. This study will help us utilize the geographic data we collect from our patients in order to better tailor these community-based interventions to our patient population.

We at the Columbus Free Clinic support this project fully. We look forward to the results of this investigation and how they may be used to help us improve the services we provide our patients. Please feel free to contact me if you wish to discuss this matter further.

Sincerely,

Haley Herman
The Ohio State University College of Medicine
M.D. Candidate, Class of 2019
Columbus Free Clinic steering committee member
Email: Haley.Herman@osumc.edu

2231 N High St. Columbus, OH 43210 □ (614) 404-8417 □ columbusfreeclinic@gmail.com

Appendix B: Institutional Review Board Initial Approval

THE OHIO STATE UNIVERSITY

Behavioral Institutional Review Board
300 Research Administration building
1960 Kenny Road
Columbus, OH 43210-1063

Phone (614) 688-8457
Fax (614) 688-0366
orrp.osu.edu

11/10/2016

Study Number: 2016B0415
Study Title: Geographic Distribution of Patients Seen by the Columbus Free Clinic

Type of Review: Initial Submission

Review Method: Expedited

Date of IRB Approval: 11/10/2016
Date of IRB Approval Expiration: 11/10/2017

Expedited category: #5

Dear Lisa Raiz,

The Ohio State Behavioral IRB **APPROVED** the above referenced research.

In addition, the following were also approved for this study:

- Waiver of Consent Process
- Full Waiver of HIPAA Research Authorization

As Principal Investigator, you are responsible for ensuring that all individuals assisting in the conduct of the study are informed of their obligations for following the IRB-approved protocol and applicable regulations, laws, and policies, including the obligation to report any problems or potential noncompliance with the requirements or determinations of the IRB. Changes to the research (e.g., recruitment procedures, advertisements, enrollment numbers, etc.) or informed consent process must be approved by the IRB before implemented, except where necessary to eliminate apparent immediate hazards to subjects.

This approval is issued under The Ohio State University's OHRP Federalwide Assurance #00006378 and is valid until the expiration date listed above. ***Without further review, IRB approval will no longer be in effect on the expiration date.*** To continue the study, a continuing review application must be approved before the expiration date to avoid a lapse in IRB approval and the need to stop all research activities. A final study report must be provided to the IRB once all research activities involving human subjects have ended.

Records relating to the research (including signed consent forms) must be retained and available for audit for at least 5 years after the study is closed. For more information, see university policies, Institutional Data and Research Data.

Human research protection program policies, procedures, and guidance can be found on the ORRP website.

Daniel Strunk, PhD, Chair
Ohio State Behavioral IRB

Association for the Accreditation of Human Research Protection Programs, Inc.
Full Accreditation

Appendix C: Institutional Review Board Amendment Approval

The Ohio State University

Behavioral Institutional Review Board
300 Research Administration building
1960 Kenny Road
Columbus, OH 43210-1063

Phone (614) 688-8457
Fax (614) 688-0366
orrp.osu.edu

01/20/2017

Study Number: 2016B0415
Study Title: Geographic Distribution of Patients Seen by the Columbus Free Clinic

Type of Review: Amendment #1

Review Method: Expedited

Request to amend the research dated December 2, 2016 (revise the protocol to reflect the the additional data collection of utilization rate data, expand the waiver of consent process to cover the new data collection, include new corresponding data coding form).

Date of IRB Approval: 01/16/2017
Date of IRB Approval Expiration: 11/10/2017

Dear Lisa Raiz,

The Ohio State Behavioral IRB **APPROVED** the above referenced research.

As Principal Investigator, you are responsible for ensuring that all individuals assisting in the conduct of the study are informed of their obligations for following the IRB-approved protocol and applicable regulations, laws, and policies, including the obligation to report any problems or potential noncompliance with the requirements or determinations of the IRB. Changes to the research (e.g., recruitment procedures, advertisements, enrollment numbers, etc.) or informed consent process must be approved by the IRB before implemented, except where necessary to eliminate apparent immediate hazards to subjects.

This approval is issued under The Ohio State University's OHRP Federalwide Assurance #00006378 and is valid until the expiration date listed above. ***Without further review, IRB approval will no longer be in effect on the expiration date.*** To continue the study, a continuing review application must be approved before the expiration date to avoid a lapse in IRB approval and the need to stop all research activities. A final study report must be provided to the IRB once all research activities involving human subjects have ended.

Records relating to the research (including signed consent forms) must be retained and available for audit for at least 5 years after the study is closed. For more information, see university policies, Institutional Data and Research Data.

Human research protection program policies, procedures, and guidance can be found on the ORRP website.

Daniel R. Strunk

Daniel Strunk, PhD, Chair
Ohio State Behavioral IRB